



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/766,308	01/27/2004	Kas Kasravi	200901335-1	1219
22879 7590 08/10/2009 HEWLETT-PACKARD COMPANY Intellectual Property Administration 3404 E. Harmony Road Mail Stop 35 FORT COLLINS, CO 80528				
EXAMINER				
SAINT CYR, LEONARD				
ART UNIT		PAPER NUMBER		
2626				
NOTIFICATION DATE		DELIVERY MODE		
08/10/2009		ELECTRONIC		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

JERRY.SHORMA@HP.COM

ipa.mail@hp.com

jessica.l.fusek@hp.com

Office Action Summary

Application No.

10/766,308

Applicant(s)

KASRAVI ET AL.

Examiner

LEONARD SAINT CYR

Art Unit

2626

Period for Reply -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 29 April 2009.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 23, 24, 26 - 31, and 33 - 40 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☒ Claim(s) 33 and 35 is/are allowed.
- 6) ☒ Claim(s) 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 23, 24, 26 - 31, 34, 36 - 40 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 01/27/04 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a). Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☐ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-848)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Response to Arguments

1. Applicant's arguments filed 04/29/09, regarding claims 1 - 3, 6, 7, 10 - 14, 16 - 21, 23, 24, and 26 - 31, and 36 - 40 have been fully considered but they are not persuasive.

Applicant argues that Gillis does not teach comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document (Amendment, page 2).

The examiner disagrees, since Gillis discloses "A vector representation of the query is calculated from the symbols (terms) in the query, and the position of this **"query vector " is located in the semantic space** obtained from the original SVD operation. **The query thus becomes a "pseudo document" in the vector space.** **The query vector is compared to the vectors of other documents in the space, and those documents which are "closest" to the query, in the semantic space, are retrieved"** (col.7, lines 59 - 66). **Comparing query vector to the vectors of other documents in the semantic space** implies comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document, **since closest documents to the query vector, in the semantic space are retrieved based on the vectors comparison.**

Applicant argues that subject matter of claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 34 and 36 - 40, as amended, is statutory subject matter (Amendment, page 2).

The examiner disagrees, since the computer, as recited in the preamble, does not perform any step of the process claims.

Claim Rejections - 35 USC § 101

2. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 34 and 36 - 40 are rejected under 35 U.S.C. 101 as not falling within one of the four statutory categories of invention. Supreme Court precedent¹ and recent Federal Circuit decisions² indicate that a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. While the instant claim(s) recite a series of steps or acts to be performed, the claim(s) neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process. Claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 34 and 36 - 40 do not identify the apparatus that accomplishes the

¹ *Diamond v. Diehr*, 450 U.S. 175, 184 (1981); *Parker v. Flook*, 437 U.S. 584, 588 n.9 (1978); *Gottschalk v. Benson*, 409 U.S. 63, 70 (1972); *Cochrane v. Deener*, 94 U.S. 780, 787-88 (1876).

² *In re Bilski*, 88 USPQ2d 1385 (Fed. Cir. 2008).

method steps like "processing the semantic vector by a digital computer" described in page 12, paragraph 36, last three lines of the specification. Thus claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 34 and 36 - 40 do not define a statutory process.

Claim Rejections - 35 USC § 102

3. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

4. Claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 23, 24, 26 - 31, and 36 - 40 are rejected under 35 U.S.C. 102(a) as being anticipated by Gillis (US Patent 6,523,026).

As per claims 1, Gillis teaches a computer-implemented method for comparing semantic content of two or more documents, comprising:

accessing a plurality of documents ("source and target domains"); performing a linguistic analysis on each document ("computing a set of vectors"; col.10, lines 9 - 17); col.11, lines 36 - 40);

defining a semantic vector for each document based on the linguistic analysis, said semantic vector having multiple components, wherein each component of said semantic vector has at least: a term included in the document or a synonym of said term; a weighting factor relating to an importance ("Selected terms within a multiterm query may be weighted, if desired, to reflect their importance to the user"), based on characteristics of the document of said term (**"The weightings of word relationships are entirely natural domain specific"**); and a frequency value relating to a number of

occurrences of said term ("computing a set of term vectors"; col.11, lines 36, and 37; col.10, lines 18 – 20; col.42, line 67- col.43, line 1; col.23, lines 35 –40); and

comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document (**"The query thus becomes a "pseudo document" in the vector space. The query vector is compared to the vectors of other documents in the space, and those documents which are "closest" to the query, in the semantic space, are retrieved"** (col.7, lines 59 – 66).

As per claim 2, Gillis further discloses that the linguistic analysis comprises sentence analysis ("sentence in the individual documents"; col.43, lines 43 – 46).

As per claim 3, Gillis further discloses that the sentence analysis comprises a syntactic analysis ("preferred stop list word include in the vectorization") and a semantic analysis ("semantic similarity"; col.39, lines 14 – 20; col.35, lines 4 – 6).

As per claim 6, Gillis further discloses that each component of the semantic vector for at least one of the documents comprises multiple dimensions ("n dimensional space"; col.39, line 63 –col.40, line 1).

As per claim 7, Gillis further discloses that each component of the semantic vector for at least one of the documents further comprises a subordinate concept value ("cable" is the subordinate concept of term "telecommunications"; col.51, lines 30 – 35).

As per claim 10, Gillis further discloses that some of the components of the semantic vector have for at least one of the documents {main term – subordinate term pairs} as their first value ("cable" and "telecommunications" are related term pairs, wherein cable is the subordinate term of telecommunications; col.51, lines 30 – 35).

As per claim 11, Gillis further discloses that the semantic vector comprises a multi-dimensional vector defined by the content of a semantic net ("n dimensional semantic space"; col.39, line 63 – col.40, line 1).

As per claim 12, Gillis further discloses that the content of the semantic net is augmented by relative weights, strengths, or frequencies of occurrence of the features within the semantic net ("frequency related weightings to term in the computation of summary vectors"; col.41, lines 40 - 46).

As per claim 14, 23, 24, 26, and 27, Gillis teaches comparing two or more documents, by:

linguistically analyzing a plurality of documents to identify at least one term group in each document, each term group comprising a main term and at least one subordinate

term semantically related to the main term ("a small subset of terms (or groups of terms such as phrases) is chosen from the source domain...computing a set of vectors"; col.10, lines 9 – 22); col.11, lines 36 – 40);

generating a semantic vector associated with each document, the semantic vector comprising a plurality of components, each component including; a term group in the document; a frequency value relating to a number of occurrences of the term group; and a weighting factor relating to an importance ("Selected terms within a multiterm query may be weighted, if desired, to reflect their importance to the user"), based on characteristics of the document of said term (**"The weightings of word relationships are entirely natural domain specific"**) of at least part of the term group ("computing a set of term vectors... Vector of terms that occurred less frequently in the training corpus are weighted more heavily in the calculation of summary vectors of search domain records"; col.11, lines 36, and 37; col.10, lines 18 – 20; col.42, line 67- col.43, line 1; col.41, lines 43 – 47; col.23, lines 35 -40); and

comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document using a defined metric (**"The query thus becomes a "pseudo document" in the vector space. The query vector is compared to the vectors of other documents in the space, and those documents which are "closest" to the query, in the semantic space, are retrieved"** (col.7, lines 59 – 66);

wherein said metric measures the semantic distance between documents as a function of at least the frequency values included in the semantic vectors for the documents ("semantically distant are individually represented at least 50 times"; col.48, lines 48 – 55; col.51, lines 30 – 35; col.41, lines 43 - 47).

As per claim 16, Gillis further discloses that the main term includes synonyms of the main term (col.11, line 8).

As per claims 17, 28, Gillis further discloses that one or more of said two or more documents are located using an autonomous software or 'bot program ("software programs"; col.10, lines 9 – 17; col.25, lines 57 – 67).

As per claims 18, and 29, Gillis further discloses automatically analyzes each document in a defined domain (source and target domains) or network by executing a series of rules and assigning an overall score to the document ("average of component values"; col.10, lines 9 – 17; col.41, line 66 –col.42, line 25).

As per claim 19, Gillis further discloses that all documents with a score above a defined threshold are linguistically analyzed ("generate term vectors and accept only records that match all the categories beyond some minimum threshold"; col.46, line 65 – col.47, line 11).

As per claims 20, and 30, Gillis further discloses that the semantic vector is a quantification of the semantic content of each document ("semantic vectors"; col.39, lines 14 – 20; col.1, lines 15 - 20).

As per claim 21, Gillis further discloses that each component has multiple dimensions ("n dimensional semantic space"; col.39, line 63 – col.40, line 1).

As per claim 31, Gillis further discloses that the output of said defined algorithm is a measure of at least one of semantic distance, semantic similarity, semantic dissimilarity, degree of patentable novelty and degree of anticipation ("semantic similarity"; col.4, lines 1 – 3).

As per claim 36, Gillis further discloses that said term comprises at least one of a word or a phrase ("a small subset of terms (or groups of terms such as phrases) is chosen from the source domain"; col.10, lines 19 – 22).

As per claim 37, Gillis further discloses that comparing the semantic vectors based on a defined algorithm (col.42, line 2).

As per claim 13, Gillis further discloses that the output of said defined algorithm is a measure of at least one of semantic distance, semantic similarity, semantic

dissimilarity, degree of patentable novelty and degree of anticipation ("semantic similarity"; col.4, lines 1 – 3).

As per claim 38, Gillis further discloses that the at least one subordinate term includes synonyms of one of the subordinate terms (col.11, line 8).

As per claim 39, Gillis further discloses that one or more of the at least one subordinate term or the main term comprises a phrase (col.10, lines 19 – 22).

As per claim 40, Gillis further discloses that the weighting factor comprises a plurality of different weighting factors and each of the different weighting factors relates to the importance of the main term or a subordinate term in the term group ("Vector of terms that occurred less frequently in the training corpus are weighted more heavily in the calculation of summary vectors of search domain records"; col.41, lines 43 - 47).

Allowable Subject Matter

5. Claims 33, and 35 are allowed over the prior art. The following is an examiner's statement of reasons for allowance:

As to claim 33, and 35, Gillis does not teach or suggest that the defined metric is one of: $\text{Sqrt} (f_1^2 + f_2^2 + f_3^2 + f_4^2 + \dots + f_{(N-1)}^2 + f_N^2) / n * 100$, wherein f is a difference in frequency of a common term between two documents and n is the number of terms those documents have in common; or $\text{Sqrt}(\sum((w-\Delta)^2))$.

$\text{Avg})/(\text{Log}(n)A^3 \cdot 1000)$, wherein $w\text{-Delta}$ is the difference in weight between two common terms, $w\text{-Avg}$ is the average weight between two common terms, and n is the number of common terms, between two documents.

Conclusion

6. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

7. Any inquiry concerning this communication or earlier communications from the examiner should be directed to LEONARD SAINT CYR whose telephone number is (571) 272-4247. The examiner can normally be reached on Mon- Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is (571)-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at (866)-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call (800)-786-9199 (IN USA OR CANADA) or (571)-272-1000.

LS

08/0309

/Michael N. Opsasnick/

Primary Examiner, Art Unit 2626